DOCUMENT RESUME

ED 278 677 TM 870 018

AUTHOR Ward, William C.; And Others
TITLE College Board Computerized Placement Tests: Validation of an Adaptive Test of Basic Skills.
INSTITUTION Educational Testing Service, Princeton, N.J.
PUB DATE Aug 86
NOTE 26p.
PUB TYPE Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS *Adaptive Testing; Basic Skills; College English; College Faculty; College Mathematics; Comparative Testing; *Computer Assisted Testing; Correlation; *Developmental Studies Programs; Grade Point Average; Higher Education; *Minimum Competency Testing; Pilot Projects; Reading Comprehension; Sex Differences; Student Placement; Student Reaction; Testing Programs; *Test Validity
IDENTIFIERS *College Board Computerized Placement Tests; Paper and Pencil Tests

ABSTRACT

The College Board Computerized Placement Tests are intended for use by two- and four-year colleges in determining whether entering students are prepared for college-level work in English, reading, and mathematics, or whether they need developmental studies. Seventeen colleges and universities participated in pilot testing of the College Board Computerized Placement Tests during 1984 and 1985. The main purposes were: (1) to familiarize the participating institutions with the adaptive testing process; (2) to determine the reactions of students, faculty and staff to the Computerized Placement Tests; and (3) to gather data relevant to the validity of the tests and their comparability to similar paper and pencil tests. Most of the 2,500 examinees were volunteers enrolled in English, reading and mathematics courses. Reactions of students and of faculty and staff were favorable. Both groups found the computer easy to use; both liked the adaptation of the tests to a student's level of skills and the ability to obtain immediate score reports. Analyses of the relations between test scores and end-of-term grades were done separately by level of course. Moderate correlations with grades were found for all categories of courses examined except for developmental mathematics courses, where the median correlations were over .50. (JAZ)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED278677

RESEARCH REPORT

COLLEGE BOARD COMPUTERIZED PLACEMENT TESTS

Validation of an Adaptive Test of Basic Skills

William C. Ward
Roberta G. Kline
Jan Flaugher

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)

✓ This document has been reproduced as received from the person or organization originating it.

☐ Minor changes have been made to improve reproduction quality.

• Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

H. Weidenmiller

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

Educational Testing Service
Princeton, New Jersey
August 1986

COLLEGE BOARD

COMPUTERIZED PLACEMENT TESTS

Validation of an Adaptive Test of Basic Skills

William C. Ward
Roberta G. Kline
Jan Flaugher

August, 1986

Copyright © 1986. Educational Testing Service. All rights reserved.

Computerized Placement Tests is a trademark
of the College Entrance Examination Board

Abstract

Seventeen two- and four-year institutions participated in pilot testing of the College Board Computerized Placement Tests during 1984 and 1985. Data were available for over 2,500 students who took one or more modules of the adaptive test; participating institutions planned to test at least 2,000 more by the end of 1985. In most institutions examinees were volunteers currently enrolled in English, reading, or mathematics courses.

Students taking part in the 1984 pilot completed a questionnaire describing their reactions, and many took a short paper and pencil test for comparison with the computerized tests. In both years of pilot activity, students' grades in relevant courses and their scores on other tests used for placement were obtained.

Reactions of students and of faculty and staff were very favorable. Both groups found the computer easy to use; both liked especially the adaptation of the test to a student's level of skills and the ability to obtain immediate score reports.

Analyses of the relations between test scores and end-of-term grades were done separately by level of course (developmental or introductory college-level). Grades in single courses constitute a very fallible criterion, confounded by a number of factors. Nonetheless, moderate correlations with grades, with median values generally in the range from .30 to .40, were found for all categories of courses examined except for developmental mathematics courses; in the latter, the median correlations were over .50. These coefficients are similar in magnitude to those found in studies with professionally constructed paper and pencil test batteries.

A few samples provided data appropriate for a comparison of relations to grades for sex and ethnic subgroups. There was no evidence of differential prediction for these groups, but, with small sample sizes and volunteer examinees, these results are far from definitive.

No specific contrasts can be made of relations to grades of the computerized tests and other placement tests given to pilot examinees; in most instances the other tests available were institution-specific, and the sample sizes were too small to permit confident generalizations about differences in the relations. Over all samples and courses for which data are available, the computerized test scores showed higher correlations with grades than did scores from paper and pencil tests in 59% of the possible contrasts.

COLLEGE BOARD

COMPUTERIZED PLACEMENT TESTS

Validation of an Adaptive Test of Basic Skills

William C. Ward
Roberta G. Kline
Jan Flaugher

The College Board Computerized Placement Tests are intended for use by two- and four-year colleges in determining whether entering students are prepared for college-level work in English, reading, and mathematics, or whether they need developmental studies in these areas. The tests consist of four modules: Reading Comprehension, Sentence Skills, Arithmetic, and Elementary Algebra.

The tests are administered on the IBM PC and compatible microcomputers, making use of Item Response Theory to provide adaptive testing. After each question is answered, the computer chooses the next question to administer based on the information it is expected to yield in light of the current estimate of the examinee's skills. Adaptive testing provides measurement as accurate as that obtained from professionally constructed conventional tests while requiring that the examinee answer fewer than half as many questions; it also provides broad range measurement, in which examinees of widely different levels of skills all receive tests appropriate in difficulty for them. Computerized testing also makes possible the immediate scoring and reporting of test results.

Seventeen colleges and universities participated in pilot testing of the College Board Computerized Placement Tests during 1984 and 1985. This testing was meant to serve several functions:

(1) To provide an opportunity for participating institutions to become familiar with the adaptive testing process and to explore its applicability to their placement testing needs;

(2) To determine the reactions of students, faculty, and staff to the Computerized Placement Tests, and to identify any difficulties they might encounter in taking or administering the tests; and

(3) To gather data relevant to the validity of the test as a measure of the basic skills needed for college work and of its comparability to paper and pencil tests intended to measure similar skills.

The purpose of this report is to summarize the results of the pilot testing, with emphasis on analyses related to the validity of the tests.

Sample

Institutions participating in the pilot testing are listed in Table 1. Twelve of these are two-year public colleges, seven with medium-sized enrollments, four with large enrollments, and one categorized as having a small enrollment. Of the five four-year institutions in the sample, three are large public universities, one is a large private university, and one is a small public university. (Information about length of program, size of enrollments and governance was taken from the College Board's The College Handbook 1984 - 85.)

In pilot activities during the spring of 1984, 426 students enrolled in three participating institutions were tested. College Board and ETS staff spent one to two weeks on campus working with faculty and staff in the data collection. In addition to the computerized tests, students completed a paper test designed to parallel one module of the adaptive test, filled out a questionnaire dealing with their reactions to the test, and in some instances took part in a small group discussion of their experiences. Students' academic records were obtained for analyses of the relations of the adaptive test scores to other tests given for placement and to grades in the English, reading, and mathematics courses in which the students were enrolled.

Subsequently, fourteen new institutions joined in a continuation of pilot testing. Most began testing in the winter or spring of 1985. In this phase, the institutions organized and administered their pilot activities without on-site assistance from College Board and ETS staff. They made their own choices as to which students would be tested, which of the four adaptive test modules would be used, and what options, such as the inclusion of institution-specific background questions, would be implemented. The only requirement for participation was an agreement to provide to project staff students' grades in relevant courses and their scores on available paper and pencil placement tests, and for pilot coordinators and their colleagues to complete questionnaires describing their reactions to the testing process.

More than 2,100 students were tested in the 1985 pilot. In most institutions test-takers were volunteers, although a few schools required intact classes or designated students to take the tests. Most students were currently enrolled in relevant courses, although one pilot site used the computerized tests as part of its regular placement process and two others used them for pre- and post-testing of students in special intervention programs for those academically at risk. Sample sizes ranged from fewer than 20 students to nearly 1,000. Excluding the samples at the extremes, the average number of students tested was about 80.

A number of participating institutions scheduled continuing pilot activity through the fall of 1985 and were expected to have tested at least 2,000 additional students by the end of 1985.

Altogether, test results for 2,539 students were available at the time this report was prepared. The numbers of students completing any one test module ranged from over 1,500 for Elementary Algebra to over 2,000 for Reading Comprehension. Sixty percent of those tested were females, and 58% identified themselves as White.

Table 1

Institutions Participating in Pilot Testing

Institution	Location	Type	Governance
Anchorage Community College	Alaska	Two-Year	Public
Bergen Community College	New Jersey	Two-Year	Public
Central Piedmont Community College	N. Carolina	Two-Year	Public
Delaware Tech. & Community College	Delaware	Two-Year	Public
Evergreen Valley College	California	Two-Year	Public
George Mason University	Virginia	Four-Year	Private
Illinois Central College	Illinois	Two-Year	Public
J. Sargeant Reynolds Community College	Virginia	Two-Year	Public
Kirkwood Community College	Iowa	Two-Year	Public
Lethbridge Community College	Alberta	Two-Year	Public
Mansfield University	Pennsylvania	Four-Year	Public
New York University	New York	Four-Year	Public
Scottsdale Community College	Arizona	Two-Year	Public
St. Louis Community College	Missouri	Two-Year	Public
Yuba College	California	Two-Year	Public
University of New Mexico	New Mexico	Four-Year	Public
University of Virginia	Virginia	Four-Year	Public

Data Collection

In the initial data collection, a standard test configuration was used in which students completed all four modules of the adaptive tests and answered a constant set of background questions. In the subsequent testing, thirteen institutions administered both verbal and mathematics tests, although not all students were given all tests; two administered only verbal tests; and one, only mathematics. Three requested a modified set of background questions and two chose the option to employ a branching rule, in which students receiving scores below a designated value in Arithmetic automatically were branched around the Elementary Algebra test.

The majority of the pilot sites tested on IBM equipment, primarily PCs but including XT's and Jr's; Compaqs and Zeniths were also used. The number of students tested at one time ranged from one to nineteen depending on the number of computers available. Institutions that had access to computer laboratories or testing centers were able to test larger numbers of students simultaneously.

Student Reactions

Students' reactions to the computerized tests, based on responses to the questionnaires administered in the 1984 pilot, are summarized in Table 2. In general, students found the computer easy to use and the tests to be of appropriate difficulty. They showed little sign of computer anxiety and an overall preference, by a substantial margin, for testing by computer rather than paper and pencil.

The questionnaires included spaces in which students wrote free-response answers to questions concerning their reactions. Answers to these questions and comments given in small group discussions indicated very positive response to several aspects of the testing situation: the ease of computer use, the absence of time limits, the challenge of a test appropriate in difficulty to their ability, and the fact that they did not have to write answers or grid ovals on answer sheets. The only frequent complaint concerned visual difficulties with the computer display: monitors were sometimes old and of low quality, and room lighting was not always designed to avoid glare. A few students also objected to being unable to return to questions previously unanswered.

Analyses were conducted to examine the relations between student characteristics and their reactions to the test. In general, there were not strong differences in reactions associated with background characteristics. For example, the few students who reported that English was their second language were less likely than native speakers to report that it was very easy to use the computer; but a majority of each group endorsed this statement.

A final set of analyses examined student reactions in relation to their test scores. Students receiving high scores on a test more often saw the test as about right in difficulty for them. High scorers tended to report being very much challenged to do their best and to have felt some time pressure; however, they were also more likely to report that the computer

Table 2

Student Reactions to the Computerized Placement Tests

How difficult was the verbal part of the computerized test for you? (that is, Reading Comprehension and Sentence Skills)

1%	much too difficult
23	somewhat too difficult
70	just about right
5	somewhat too easy
1	much too easy

How difficult was the mathematics part of the computerized test for you? (that is, Arithmetic and Elementary Algebra)

9%	much too difficult
37	somewhat too difficult
41	just about right
12	somewhat too easy
1	much too easy

Did you feel challenged to do as well as you could on the test?

6%	not at all
52	somewhat
42	very much

In comparison with a paper and pencil test, did you find the computerized test

48%	more challenging than a paper and pencil test
39	about the same
13	less challenging than a paper and pencil test

Were you nervous while taking the test?

61%	not at all
34	somewhat
5	very much

Did nervousness while taking the test prevent you from doing your best?

3%	yes, definitely
17	yes, somewhat
44	probably not
36	definitely not

Which would make you more nervous, a paper and pencil test or a computer-administered test?

43%	a paper and pencil test
11	a computer-administered test
46	they'd be about the same

Did you feel any pressure to complete the test quickly?

3%	yes, a great deal of pressure
38	yes, some pressure
59	no

Table 2 (continued)

How easy was it to use the computer to take the test?

79%	very easy
19	reasonably easy
2	somewhat difficult
0	very difficult

Were the directions given by the computer clear and easy to follow?

87%	yes, very clear
12	reasonably clear
1	no, not very clear

Did you have any difficult reading the material presented on the computer screen?

23%	yes
77	no

In comparison with a paper and pencil test of the same length, did you find the computer test

8%	more tiring than a paper and pencil test
32	about the same
60	less tiring than a paper and pencil test

Which would you prefer to take?

74%	a computer-administered test
10	a paper and pencil test
16	no preference

was very easy to use and the directions were very clear, and that nervousness did not prevent them from doing their best. They did not differ from low scorers in their prior experience in using a computer, but were more likely to report having programmed one.

Reactions of Faculty and Staff

Questionnaires were sent to each of the institutions participating in the 1985 pilot testing. Eight testing coordinators completed an extensive questionnaire which sought their reactions and descriptions of the conditions under which testing took place; 30 additional faculty and staff members who had had an opportunity to try out the system completed a briefer form dealing primarily with their own experience and reactions to the test.

Both groups responded favorably to the test. They reported that they and their students found the directions clear and the computer easy to use; a strong majority believed that students would prefer a computerized test to a paper and pencil one. Among the aspects of the test they liked best were its adaptation to a student's level of skills and the immediate scoring and reporting of results. Among those liked least were the inability to return to change the answer to a previously administered question, an option not appropriate in adaptive testing, and the lack of diagnostic information. The latter need will be addressed by the diagnostic test of basic skills currently under development by The College Board and ETS.

Descriptive Results of Pilot Testing

The mean scores for all students tested in the 1984 and 1985 pilots are presented in Table 3. For Reading Comprehension, the mean score corresponds to the 46th percentile in a normative sample, comprised of a representative sample of students taking a statewide paper and pencil placement battery. For Sentence Skills, Arithmetic, and Elementary Algebra, respectively, the means correspond to the 33rd, 42nd, and 42nd percentiles.

Also shown in Table 3 are the means for the most able and least able samples of students tested. On Reading Comprehension, for example, the most able sample tested at any participating institution had a mean score corresponding to the 68th percentile in a normative sample, while the least skilled had a mean corresponding to the 19th percentile. In every sample the typical student tested fell within the range of levels of skills for which the tests were intended to provide accurate measurement.

Table 4 displays median correlations among the test modules across all samples tested. As would be expected, the two mathematics tests and the two verbal tests have the highest intercorrelations. All the tests, however, are moderately correlated with one another.

Table 3

Mean Scores on the Computerized Placement Tests

		Reading Comprehension	Sentence Skills	Arithmetic	Elementary Algebra
All Students	M	75.86	82.95	67.21	55.47
	Percentile	46	33	42	42
	N	2021	1934	1686	1547
High	M	90.89	102.10	90.52	98.59
	Percentile	68	67	72	84
	N	70	62	62	62
Low	M	49.48	55.55	54.34	34.41
	Percentile	19	11	29	20
	N	76	76	125	42

Table 4

Median Correlations Among Scores on the Computerized Placement Tests

	Sentence Skills	Arithmetic	Elementary Algebra
Reading Comprehension	.54	.49	.34
Sentence Skills		.44	.36
Arithmetic			.66

Relations of the Computerized Tests to Course Grades

Several kinds of evidence can be employed to demonstrate the validity of placement tests. The most frequently used, and the one employed here, is the study of relations between the test scores and students' grades in relevant courses. To the degree that the tests measure skills important for success in those courses, they should be correlated with course grades.

Analyses of relations between test scores and end-of-term grades were done separately by level of course (developmental or introductory college-level); combining students across levels would have introduced a source of considerable extraneous variation into the relations. Correlations are reported whenever a group of 20 or more students was available. For the analysis, grades were coded using the nine-point scale shown in Table 5.

Average results across all the samples available for analysis show very respectable relations between grades and test scores. The results are displayed in Tables 6-10 and summarized below.

It should be noted that there is sometimes a severe limit on the magnitude of correlations that can be expected from these analyses. Grades in single courses are not a very reliable criterion under the best of circumstances. Moreover, in many instances the analysis required that results for students enrolled in a number of different courses be combined into a single group so that a sufficient number of cases would be available for analysis; there may, for example, have been as many as half a dozen courses all classified as developmental English. To the degree that different courses entail different degrees of rigor, either because of the course objectives or because of differences in instructors' grading practices, the correlations will be adversely affected.

Further, in several instances the distribution of grades was so restricted in range that appreciable correlations would be very unlikely. For example, in one institution the grades given in developmental reading ranged only from Satisfactory to Unsatisfactory, and more than 70% of all those graded received a grade of Satisfactory. Finally, to the degree that students assigned to a given level of coursework have been appropriately placed for their level of skills, the computerized test scores will also be restricted in range.

Developmental English

Eight institutions tested a sufficient number of students enrolled in developmental English to permit correlational analysis of their data. Several provided more than one sample of students; for example, an institution participating in both years of pilot testing, or one testing several groups (such as regular enrollees and students in an enrichment program), could contribute two samples to the analysis. In Table 6 and subsequent tables each institution is identified by a letter and distinct samples from that institution are presented as separate entries.

Table 5

Coding of Grades

Grade	Numerical Code
A+, A	9
A-	8
B+	7
B	6
B-	5
C+	4
C	3
C-	2
D+, D, D-	1
F	0
WITHDRAWN	-

Table 6

Correlations of Scores on the Computerized Placement Tests with End-of-Term Grades in Developmental English Courses

Institution	Type	Reading Comprehension N	Reading Comprehension r	Sentence Skills N	Sentence Skills r
A	Two-Year Public	68	.40	68	.52
B	Two-Year Public	48	.64	47	.49
B	Two-Year Public	23	.32	--	--
C	Two-Year Public	44	.39	44	.16
C	Two-Year Public	62	.25	62	.26
D	Two-Year Public	66	-.05	66	.10
E	Two-Year Public	39	.09	39	.39
F	Four-Year Public	105	.39	83	.42
G	Four-Year Public	21	-.06	21	.49
G	Four-Year Public	26	.05	26	.14
H	Four-Year Public	63	.52	62	.60
		---	---	---	---
	LOW		-.06		.10
	HIGH		.64		.60
	MEDIAN		.32		.40

Note: N is the number of cases on which the correlation is based; r is the correlation coefficient.

Over all samples, the median correlation of Reading Comprehension with grades was .32, while that of Sentence Skills with grades was .40. These median values are similar to those reported for paper and pencil tests used for placement. For example, for a sample of four colleges reporting relations of the Descriptive Tests of Language Skills to grades in developmental writing courses, the median correlation of the DTLS Reading Comprehension score with grades was .31, while that of the DTLS Sentence Structure score with grades was .28 (Guide to the Use of the Descriptive Tests of Language Skills, 1985).

College English

Results for college English courses are presented in Table 7. Four samples from three institutions were available for analysis. The correlations varied substantially from sample to sample; overall the median correlation for Reading Comprehension was .30, while that for Sentence Skills was also .30. Comparative results for the DTLS tests show a median correlation of .43 between DTLS Reading Comprehension and grades in college-level writing courses, and of .38 between DTLS Sentence Structure and those grades. For the College Board Test of Standard Written English, the median correlation with Freshman English grade for institutions using the Validity Study Service was .38 (Donlon, 1984); the median correlation in six samples analyzed by Breland (1977) was .28.

Developmental Reading

Table 8 presents correlations for reading courses for samples from six institutions. Again there is substantial variation from sample to sample; overall the median correlation was .29 for the relation of grades to Reading Comprehension, and .35 for that of grades to Sentence Skills.

Developmental Mathematics

In Table 9 are displayed correlations for developmental mathematics courses. Seven samples from six institutions were available for analysis. Most samples showed substantial relations. The median correlation with grades was .52 for both the Arithmetic and the Elementary Algebra score.

These relations are similar in magnitude to those found by Bridgeman and Kline (1980) in a study of the validity of the Descriptive Tests of Mathematics Skills. The DTMS Arithmetic Skills test given at the beginning of the semester had a median correlation of .52 with end-of-term grades in developmental mathematics courses (21 samples), while the DTMS Elementary Algebra Skills test had a median correlation of .43 with those grades (18 samples).

Table 7

Correlations of Scores on the Computerized Placement Tests with End-of-Term Grades in College English Courses

Institution	Type	Reading Comprehension N	Reading Comprehension r	Sentence Skills N	Sentence Skills r
A	Two-Year Public	41	.53	41	.47
B	Four-Year Public	47	.31	47	.47
B	Four-Year Public	29	.14	29	.12
C	Four-Year Private	32	.29	32	.14
		---	---	---	---
	LOW		.14		.12
	HIGH		.53		.47
	MEDIAN		.30		.30

Note: N is the number of cases on which the correlation is based; r is the correlation coefficient.

Table 8

Correlations of Scores on the Computerized Placement Tests with End-of-Term Grades in Reading Courses

Institution	Type	Reading Comprehension N	Reading Comprehension r	Sentence Skills N	Sentence Skills r
A	Two-Year Public	33	.74	33	.69
B	Two-Year Public	38	.41	38	.44
C	Two-Year Public	65	-.11	65	-.01
D	Two-Year Public	43	.03	--	--
E	Two-Year Public	43	.25	42	.25
F	Four-Year Public	43	.33	43	.35
		---	---	---	---
	LOW		-.11		-.01
	HIGH		.74		.69
	MEDIAN		.29		.35

Note: N is the number of cases on which the correlation is based; r is the correlation coefficient.

Table 9

Correlations of Scores on the Computerized Placement Tests with End-of-Term Grade in Developmental Mathematics Courses

Institution	Type	Arithmetic N	Arithmetic r	Elementary Algebra N	Elementary Algebra r
A	Two-Year Public	39	.64	38	.64
B	Two-Year Public	52	.52	52	.58
B	Two-Year Public	33	.35	--	--
C	Two-Year Public	34	.20	34	.01
D	Four-Year Public	28	.54	28	.56
E	Four-Year Public	63	.53	63	.41
F	Four-Year Public	61	.44	61	.49
		---	---	---	---
	LOW		.20		.01
	HIGH		.64		.64
	MEDIAN		.52		.52

Note: N is the number of cases on which the correlation is based; r is the correlation coefficient.

College Mathematics

College mathematics course results were available for three institutions. The median correlations with grades (Table 10) were .42 for Arithmetic and .31 for Elementary Algebra. Bridgeman & Kline reported a median correlation of .51 for the relation between the DTMS Elementary Algebra Skills test and grades in seven college algebra courses. Further data for comparison are available for 29 colleges that through the Validity Study Service have studied the relationship between the mathematical section of the SAT and grades in a freshman mathematics course. For those institutions the median correlation was .33.

In the results summarized above, there is often great variability in the magnitude of the correlations from one sample to another. Median results across samples are much more consistent. In general, the computerized test scores showed average correlations in the range .30 to .40 with grades in English, reading, and mathematics courses; the one exception was that of the relations of test scores to developmental mathematics grades, where the median correlations were over .50. These coefficients are similar in magnitude to those found in studies with standardized paper and pencil batteries, and are about as high as can reasonably be expected given the variety of extrinsic factors than operate to attenuate the relationship.

It should be noted that the relations described above arise in courses for which placement was not made on the basis of the computerized tests. If the tests were employed in placement, the relations of test scores with grades would not be expected to be as great, since the range of scores for students placed in a given level of coursework would be restricted; moreover, to the degree that students were appropriately placed, the grades received in the course would also be more homogeneous.

In addition to the results for intact samples, relations with grades were examined for sex and ethnic subgroups in each sample in which contrasting groups of at least 20 cases were available.

Possible subgroup differences are of interest given the frequently expressed concern, largely in the absence of relevant data, that computer anxiety or lack of experience might differentially affect performance of some subgroups in computerized testing.

Ten instances were identified in which males and females could be contrasted: six samples in developmental English courses, three in developmental mathematics, and one in college mathematics. Only one sample yielded significant differences; males enrolled in developmental English and mathematics courses in one institution showed higher correlations of test scores with grades than did females, probably because the computerized test scores for males were substantially more heterogenous. No overall tendency for higher correlations for one or the other group was found.

Three instances were identified in which Whites and Nonwhites could be contrasted: two in developmental English courses and one in developmental mathematics. In only one sample were there significant differences; for students enrolled in a developmental English course, the correlation of Reading Comprehension with grades was greater for Nonwhites than for Whites. Again no overall trend was found.

Table 10

Correlations of Scores on the Computerized Placement Tests with End-of-Term Grades in College Mathematics Courses

Institution	Type	Arithmetic N	Arithmetic r	Elementary Algebra N	Elementary Algebra r
A	Two-Year Public	49	.41	48	.19
B	Two-Year Public	39	.57	39	.38
C	Four-Year Public	31	.42	31	.31
		---	---	---	---
	LOW		.41		.19
	HIGH		.57		.38
	MEDIAN		.42		.31

Note: N is the number of cases on which the correlation is based; r is the correlation coefficient.

These comparisons are based on small samples, and there is little reason for confidence that the subgroups, depending largely on volunteer examinees, are representative ones; thus they are far from definitive. The strongest statement that can be made from them is that there is no evidence for systematic difference in prediction of grades based on ethnic or sex group membership.

Relations of the Computerized Tests to Other Placement Tests

The most systematic basis for comparing the computerized tests with paper and pencil tests is provided by data collected as part of the 1984 pilot. The majority of students participating in the pilot were asked to complete a short paper and pencil test constructed from the same pool of items as a module of the adaptive tests but sharing no questions with it. Two hundred thirty-four took a reading comprehension test, while 88 completed an elementary algebra test. The results indicated that there were no systematic differences in level of students' performance associated with the mode of testing, and that the relations between corresponding tests, when corrected for test reliability, were nearly perfect.

Institutions participating in the 1985 pilot were asked to provide scores on any available placement tests completed by their students. As shown in Table 11, six schools provided eight samples in which relations of Reading Comprehension to other verbal tests could be examined. The correlations ranged from a low of .17 to a high of .72, with a median value of .50. Similar analyses for the remaining test modules indicated median correlations of .55 for Sentence Skills, .54 for Arithmetic, and .42 for Elementary Algebra.

In most instances, the tests used were institution-specific and the sample sizes were too small to support generalizations about the "true" relations on the basis of one or two samples. Therefore no more detailed presentation of the results is provided.

For similar reasons no specific contrasts will be reported between the computerized tests and paper and pencil tests in predicting the grades of pilot participants. Over all samples and all courses, there were 61 instances in which the magnitude of correlations could be contrasted. One institution, for example, provided scores on three locally administered mathematics tests and both computerized mathematics tests for students enrolled in developmental mathematics; six of the 61 possible contrasts (two computerized tests times three paper tests) arise from its data. Altogether, the computerized test scores showed higher correlations with grades than did the paper test scores in 59% of the contrasts made.

Summary

Seventeen two- and four-year institutions participated in pilot testing of the College Board Computerized Placement Tests during 1984 and 1985. Data were available for over 2,500 students who took one or more modules of the

Table 11

Correlations of Scores on the Computerized Placement Tests with Other Tests Given for Placement

	Number of Institutions	Number of Samples	Low r	High r	Median r
Reading Comprehension	6	8	.17	.72	.50
Sentence Skills	5	7	.31	.78	.55
Arithmetic	4	6	.09	.76	.54
Elementary Algebra	4	6	.04	.77	.42

adaptive tests. Students participating in the 1984 pilot completed a questionnaire describing their reactions, and many took a short paper and pencil test for comparison with the computerized test. In both years of pilot activity, students' grades in relevant courses and their scores on other tests used for placement were obtained.

Reactions of students and of faculty and staff were very favorable. Both groups found the computer easy to use; both liked especially the adaptation of the test to a student's level of skills and the capability to obtain immediate score reports.

The pilot provided an opportunity to evaluate the effectiveness of the computerized test in predicting grades in English, reading, and mathematics. Grades in single courses constitute a very fallible criterion, confounded by a number of factors. Nonetheless, moderate correlations with grades, with median values generally in the range from .30 to .40, were found for all categories of courses examined except for developmental mathematics courses; in the latter, the median correlations were over .50. These coefficients are similar in magnitude to those found in studies with professionally constructed paper and pencil test batteries, indicating that the computerized tests have validity comparable to such batteries.

Continuing studies of the validity of the tests will be conducted through the College Board's Placement Research Service. As institutions begin using the tests to place students, several kinds of evidence regarding validity will accumulate: further studies of correlations of test scores with grades, similar to those reported here, and studies of the accuracy of placement decisions for students at varying levels of skills as shown by their test scores.

References

Breland, H. M. (1977). A Study of college English placement and the Test of Standard Written English. College Board Research and Development Report 76-77, No. 4. Princeton, NJ: Educational Testing Service.

Bridgeman, B., & Kline, R. G. (1980). Validation of the Descriptive Tests of Mathematics Skills. College Board Research and Development Report 79-80, No. 2. Princeton, NJ: Educational Testing Service.

College Board (1984). The college handbook 1984 - 85. New York: College Entrance Examination Board.

College Board (1985). Guide to the use of the Descriptive Tests of Language Skills. New York: College Entrance Examination Board.

Donlon, T. D., ed. (1984). The College Board technical handbook for the Scholastic Aptitude Test and Achievement Tests. New York: College Entrance Examination Board.